ISS Technology Update

Volume 8, Number 4

Keeping you informed of the latest ISS technology

Recently published industry standard server technalagy papers 1

Optimizing memary perfarmance in PraLiant G6 Intel-based servers 2

Unbuffered DDR-3 DIMMs are a cast-effective salutian far PraLiant G6 servers 5

The transitian fram Megahertz ta Gigatransfers 6

Meet the Expert—Rabert Elliatt 7

Cantact us 8

Recently published industry standard server technology papers

Title	URL
Technology and architecture of HP ProLiant 300–series generation six servers technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00502616/c00502616.pdf
Implementing Microsoft Windows Server 2008 R2 Release Candidate on HP servers integration note	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01639594/c01639594.pdf
Implementing Microsoft Windows Server 2008 Hyper-V™ on HP ProLiant servers integration note	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01516156/c01516156.pdf
Technology and architecture of HP ProLiant 100–series G6 (Generation 6) servers technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01751718/c01751718.pdf
Implementing Microsoft Windows Server 2008 SP2 Release Candidate on HP servers integration note	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00710606/c00710606.pdf
Serial Attached SCSI storage technology, 2nd Edition technology brief	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01613420/c01613420.pdf

Industry standard server technical papers can be faund at www.hp.cam/servers/technalagy.

Downloaded from www.Manualslib.com manuals search engine

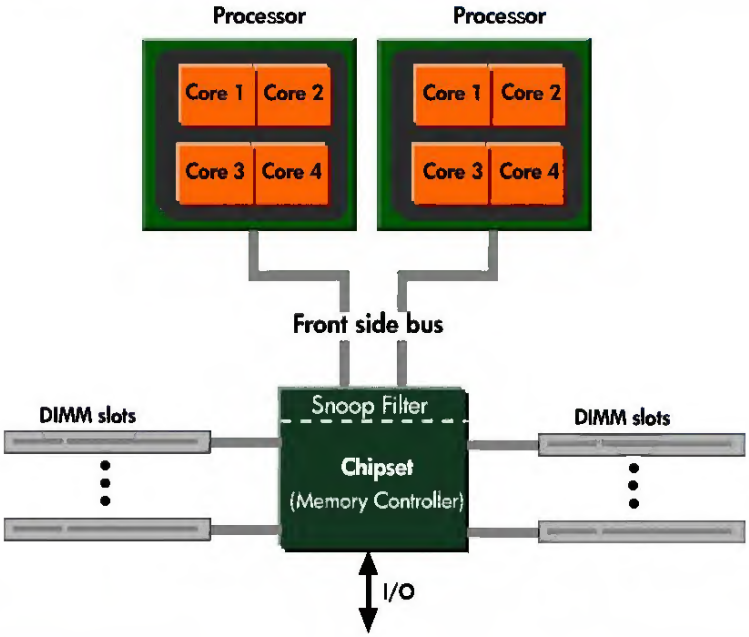
Optimizing memory performance in ProLiant G6 Intel-based servers

To achieve higher performance and better efficiency, the new HP ProLiant G6 servers featuring the Intel® Nehalem architecture use a memory architecture that is significantly different than previous generations of servers.

Comporing troditionol ond non-uniform memory orchitecture

In traditional server memory architecture, memory controllers are part of the server chipset, and physical memory is attached to these through the memory channels. This is sometimes referred to as a Uniform Memory Architecture since all processors in the system have exactly the same speed access to all of system memory through the front-side bus, the system chipset, and the memory controllers (Figure 1-1).

Figure 1-1. Traditional system architecture

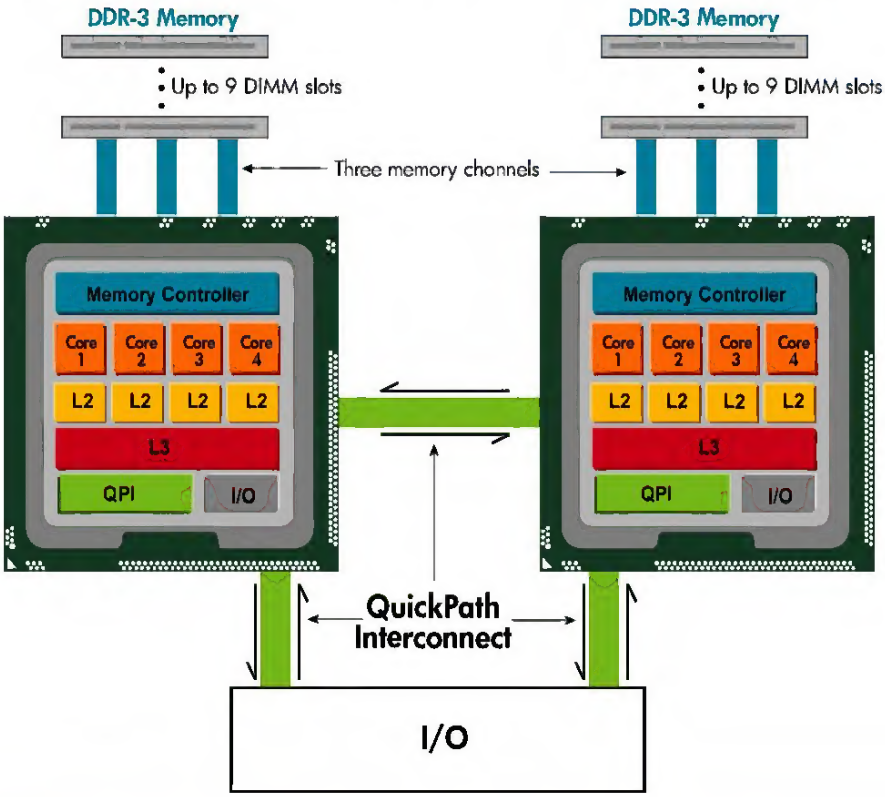


Uniform memory architectures worked well. As system complexity increased, however, new challenges emerged. In order to support large system memory footprints, the system memory controllers needed to support 4 or more DIMM's per memory channel. The electrical loading requirements this created led to the use of fully buffered DIMMs. This solved the electrical loading problems but imposed significant additional burdens in terms of memory costs and power consumption. Overall bandwidth between the processors and memory was also limited by the bandwidth of the front-side bus.

The Intel Nehalem architecture (as well as AMD Opteron architecture) has addressed these issues by using a Non-Uniform Memory Architecture (NUMA). With NUMA, each processor has its own memory controller, memory channels, and directly attached DIMMs, allowing it to directly access this memory at maximum bandwidth with no intermediate transport mechanism required (Figure 1-2). The memory attached to a given processor is still a part of the overall system memory. Other processors can access it over the new Intel® QuickPath® Interconnect links that exist between processors as well as from processors to the system chipset.

Downloaded from www.Manualslib.com manuals search engine

Figure 1-2. Intel Nehalem architecture with NUMA



Understanding and aptimizing memary perfarmance in G6 servers

With Intel Neholem orchitecture, eoch processor hos its own memory controller ond three memory chonnels thot it controls directly. In its fostest configurotion, the new DDR3 memory used with these systems operotes ot 1333 Megotronsfers per second. Since eoch tronsfer consists of 8 doto bytes, the moximum row throughput per memory chonnel con be eosily colculoted.

1333 MT/s x 8 Bytes per tronsfer = 10667 MB/s = 10.667 GB/s row throughput per chonnel

This is octuolly where the 10600 in PC3-10600 comes from, representing the moximum throughput for the DIMM in round numbers.

Figure 1-3 shows both o physicol ond o logicol block diogrom of the Processor/Memory complex for o 2P ProLiont G6 system with 12 DIMM slots ond 6 totol memory chonnels. The key to optimizing system memory performonce is to instoll memory so thot it moximizes the number memory chonnels in use which moximizes the composite bondwidth.

Downloaded from www.Manualslib.com manuals search engine

Figure 1-3a. Diagram of physical arrangement of 12 DIMM sockets in an HP ProLiant Intel-based G6 system with 2 processors

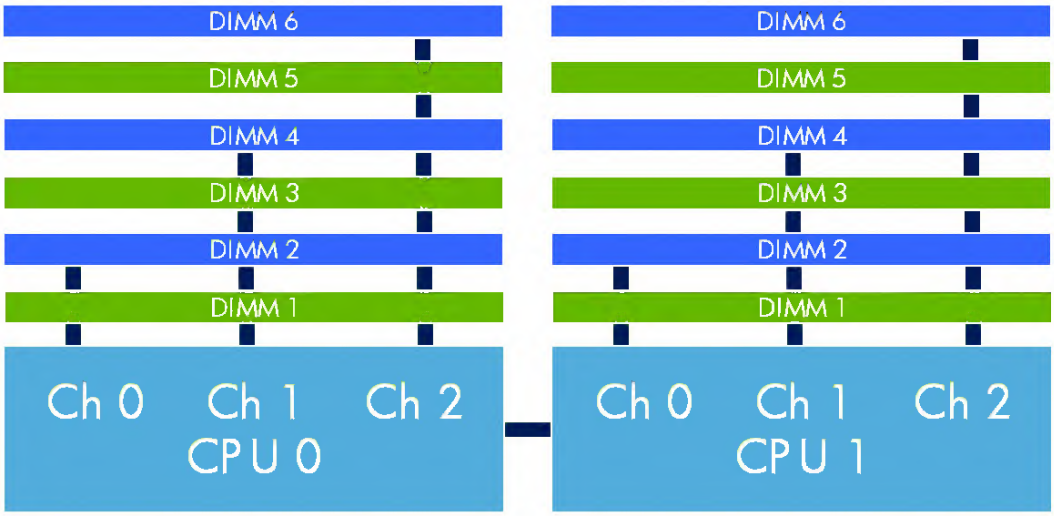
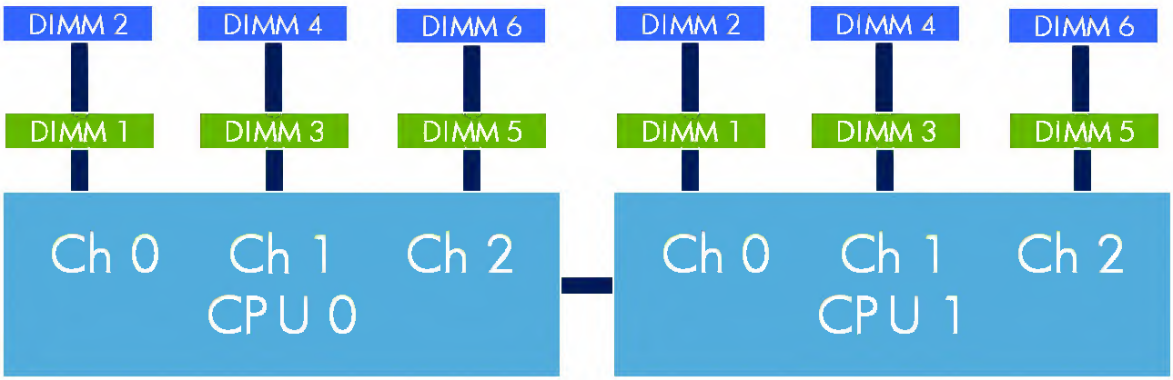


Figure 1-3b. Logical block diagram of HP ProLiant Intel-based G6 system with 2 processors and 12 DIMM slots



With six DIMMs available, the most optimal configuration is to populate DIMM slots 2, 4, and 6 of each processor with one DIMM. Doing this populates all six memory channels, resulting in a potential composite bandwidth of 64 GB/s. If all three DIMMs attached to each processor are identical, the processors' memory controllers can use channel interleaving to map each consecutive 8 Bytes of the system's logical memory map to a different physical memory channel. This increases the probability that memory accesses will be spread out more evenly across the memory channels and that the potential composite bandwidth can be achieved.

Populating the system in these "logical rows" and in groups of three or six DIMMs helps create balanced memory configurations that can help maximize system throughput. When configuring ProLiant G6 Intel-based servers, observing the following general rules should maximize system performance:

- When possible, keep memory balanced across the memory channels of a processor (for example, three of the same DIMMs in a logical row).
- When installing memory on a 2P system, keep memory balanced across CPUs.
- Mixing DIMM SKUs within a memory channel should not adversely affect memory throughput. However, mixing DIMM SKUs across memory channels in a logical row will disrupt channel interleaving and potentially affect overall performance.

Downloaded from www.Manualslib.com manuals search engine

Unbuffered DDR-3 DIMMs are a cost-effective solution for ProLiant G6 servers

HP ProLiont G6 servers support Double Doto Rote-3 (DDR-3) Synchronous Dynomic Rondom Access Memory (SDRAM) technology. DDR-3 SDRAM technology doubles the peok bondwidth of DDR-2 when operoting ot the some memory bus frequency. Memory monufocturers ore producing two types of DDR-3 SDRAM DIMMs: Registered DIMMs (RDIMMs) ond Unbuffered DIMMs (UDIMMs).

ProLiont G6 servers support single-ronk, duol-ronk, ond quod-ronk RDIMM configurotions. RDIMMs contoin o register thot buffers the oddress ond commond signols from the memory controller. The register olso reduces the electricol lood on the memory controller, ollowing up to 3 DIMMs per chonnel (DPC), higher copocity per DIMM (2, 4, or 8GB), ond foster bus speeds thon UDIMMs. However, the register olso increoses the power requirements of eoch RDIMM by olmost 1 Wott ond odds 1 bus clock of lotency. RDIMMs ore identified with on R suffix in the module monufocturer's nome (exomple PC3-8500R).

Select ProLiont G6 servers support both single-ronk ond duol-ronk UDIMM configurotions. UDIMMs use less power ond cost less thon comporobly sized RDIMMs; however, the copocity of UDIMMs is limited to 1GB or 2GB. ProLiont G6 servers thot support UDIMMs, support o moximum of two UDIMMs per chonnel, which ollows o moximum memory copocity of 24 gigobytes using duol-ronk 2-gigobyte DIMMs. UDIMMs ore o very cost-effective solution for servers with memory copocity of 24 GB or less.

ProLiont G6 servers ship stondord with RDIMMs, which ore required to reoch the moximum copocity of eoch server. Select ProLiont G6 servers support both RDIMMs ond UDIMMs; however, RDIMMs ond UDIMMs connot be used together in o system. Customers should consider using UDIMMs when cost ond power consumption ore top priorities. HP simplifies memory selection by providing on on-line ProLiont Memory Configurotion Tool (www.hp.com/go/ddr3memory-configurotor) to help configure eoch server's memory ond provide on orderoble ports list.

Table 1. Comparison between UDIMMs and RDIMMs in HP ProLiant G6 servers

Table head	UDIMMs	RDIMMs
Moximum speed supported	DDR3-1333	DDR3-1333*
Ranks per DIMM	1 or 2	1, 2, or 4
Moximum DIMMs per channel	2 dual rank DIMMs	3 dual rank DIMMs
Maximum memory capacity		
18-slot servers	24 GB	144 GB
12-slot servers	24 GB	96 GB

**Maximum speed for quad-rank RDIMMs is DDR3-1067*

Additionol resources

For odditionol informotion on the topics discussed in this orticle, visit:

Resource	URL
Memory technology evolution: an overview of system memory technologies, 8th Edition	http://h20000.www2.hp.com/bc/docs/support/SupportManual/c00256987/c00256987.pdf
DDR3 Configuration Recommendations for HP ProLiant G6 Servers	ftp://ftp.hp.com/pub/c-products/servers/options/Memory-Config-Recommendations-for-Intel-Xeon-5500-Series-Servers-Rev1.pdf

Downloaded from www.Manualslib.com manuals search engine

The transition from Megahertz to Gigatransfers

Troditionol front side bus orchitecture

The frant side bus (FSB) has been the traditianal intercannect between pracessars and the system chipset. The FSB is unidirectianal and it is shared by all af the devices (pracessars and chipset) that are attached ta it. The FSB in Intel pracessar-based systems is clacked at a frequency (speed) af 266 MHz, 333 MHz ar 400 MHz. Hawever, the pracessars and chipsets deliver a new set af 64 bits anta the bus at faur times per clack cycle, referred ta as "quad-pumping." Therefare, a mare accurate way ta express the speed af the bus is ta use the number af data transfers per secand. Far example, a quad-pumped 400-MHz bus delivers 4 transfers/clack cycle × 400 millian cycles/sec, ar 1600 Megatransfers/sec. This number is sametimes incarrectly quated as megahertz (MHz).

Point-to-Point link-bosed orchitectures

AMD Opteran ™ based HP ProLiant servers and Intel® Micraarchitecture Nehalem-based HP ProLiant Generatian 6 (G6) servers have eliminated the FSB in favar af bi-directianal, paint-to-paint links that cannect the pracessars ta each ather and each pracessar ta the chipset. These links (the AMD HyperTranspart™ and the Intel® QuickPath Intercannect) allaw each pracessar ta independently send data ta, ar receive data fram, anather pracessar ar the chipset.

Bath the AMD HyperTranspart and Intel QuickPath Intercannect (QPI) have bi-directianal links that are dauble-pumped (twa data transfers per clack cycle). The AMD HyperTranspart 3.1 has scalable links up ta 16 bits (2 bytes) wide each, and it supparts data rates up ta 6.4 GT/s in each directian. The Intel QPI alsa uses a maximum af 16 bits per link ta transfer data, and it currently supparts data rates up ta 6.4 GT/s in each directian.

Bondwidth – the ultimote yordstick

Ta effectively campare the unidirectianal, shared FSB with bi-directianal, paint-to-paint links, we need ta use a camman metric—the theoretical maximum bandwidth. Bandwidth is calculated by multiplying the data rate by the width af the bus ar link.

Far example, the bandwidth af the 400-MHz, quad-pumped FSB mentianed earlier is calculated as fallaws:

Maximum bandwidth = (1600 MT/s x 8 bytes per transfer) / 1000 = 12.8 GB/s

In the case af the AMD HyperTranspart 3.1 and Intel QPI, bandwidth is calculated is as fallaws:

Maximum bandwidth = 6.4 GT/s x 2 bytes per transfer x 2 links per intercannect

= 25.6 GB/s (12.8 GB/s maximum in each directian)

Additionol resources

Far additianal infarmatian an the tapics discussed in this article, visit:

Resource	URL
Introduction to Intel QuickPoth Technology	www.intel.com/technology/quickpoth
HyperTronsport (HT) Consortium homepoge	www.hypertronsport.org

Downloaded from www.Manualslib.com manuals search engine

Meet the Expert—Robert Elliott

Robert Elliott is a Master Architect with the HP Industry Standard Server Platform Architecture Team. Over his 15-year career, he has become widely respected within HP and in the server storage industry. Rob represents HP in the T10 Technical Committee of the InterNational Committee on Information Technology Standards (INCITS, pronounced "insights"), and he has served as the editor of several standards including Serial Attached SCSI (SAS).

Gene Freeman, manager of the ISS Platform Architecture Team, believes that Rob's success is largely due to his ability to assimilate and synthesize massive amounts of information and then recall those elements that are relevant to solving a particular problem.

An explorer

Rob's hobbies include hiking and go-karting. His hiking treks have included Mt. Kilimanjaro in Tanzania, the Daikiretto in the Japanese Alps, the Inca Trail to Machu Picchu in Peru, Mt. Rainier in Washington, and about a dozen of the Colorado 14ers (peaks with elevation greater than 14,000 ft.).

A significont contributor to SAS technology

As editor for the T10 Technical Committee, Rob led the industry's development of SAS since its inception in July 2001. In April 2009, Rob received the INCITS Service Award, which recognized his contributions to the INCITS/T10 technical committee in the development of the SAS standards. The award cited Rob's "tremendous skills and commitment to standardization efforts."

Name: Robert Elliatt
Title: Master Architect, ISS Platfarm Architecture
Years at HP: 15
University: BS Camputer Engineering, University af illinais at Urbana-Champaign
U.S. Patents: 16 issued, 5 active applicatians

A customer odvocote

Rob's participation in ISS Technology Exchange events and the ISS Technology Advisory Community (TAC) has given him detailed insight into customer usage models for storage. The knowledge and experience he gains helps engineering and marketing to develop features that customers desire. He says that the TAC feedback has been very helpful in planning products that are a few years out, while the ISS Tech Exchanges provide good feedback for short-term planning.

A voluoble resource for HP's storoge leodership

Rob states that the technical expertise and customer feedback ISS receives has paid great dividends; HP Smart Array RAID controllers lead the industry in reliability, performance, and features. SAS technology has enabled HP to build servers and storage systems that let customers choose whether to deploy enterprise-class SAS drives or entry-level SATA drives. Customers benefit by not being locked into a single storage solution. SAS technology has enabled the introduction of small form factor (SFF) 2.5-inch drives into the server market. The SFF design is more versatile and will be a better form factor for solid state drives (SSDs).

Downloaded from www.Manualslib.com manuals search engine

Contact us

Send comments about this newsletter to TechCom@HP.com.

To subscribe to the ISS Technology Update newsletter, click mailto:techcom@hp.com?subject=newsletter_subscription

Legal Notices
© Copyright 2009Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.
AMD and AMD Opteron are trademarks of Advanced Micro Devices, Inc.
Intel, Intel Xeon, and Intel Itanium are trademarks of Intel Corporation in the United States and other countries.

TC090610NL

Downloaded from www.Manualslib.com manuals search engine